DOCUMENT RESUME

ED 305 393 TM 012 947

AUTHOR Fox, Kathleen V.
TITLE A Modified Mastery System and a Modified Norm-Referenced Grading System Compared.
PUB DATE Feb 88
NOTE 17p.; Paper presented at the Annual Meeting of the Eastern Educational Research Association (Miami Beach, FL, February 24-27, 1988).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; *Grading; Higher Education; *Mastery Tests; *Norm Referenced Tests; *Scoring Formulas; *Undergraduate Students
IDENTIFIERS *Modified Mastery System; *Modified Norm Referenced System

ABSTRACT

A comparison was made between scores and grades of college students taking a development and learning course, using either a modified mastery grading system (MMGS) or a modified norm-referenced grading system (MNRGS). Under the MMGS, students could take each unit exam up to three times to meet minimum or higher criteria levels. Under the MNRGS, students took only one test per unit and were assigned grades by comparing their score on the unit (or course average) to the mean of the upper 10% of the distribution of scores (or averages). Although there was improvement in the MMGS group scores and a greater distribution of grades under the MNRGS, repeated testing under the MMGS did not yield significantly higher averages on unit tests or final course averages than those found in the MNRGS group. Four data tables are included. (Author/TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED305393

A MODIFIED MASTERY SYSTEM

and

A MODIFIED NORM-REFERENCED GRADING SYSTEM

COMPARED

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Kathleen V. Fox

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Kathleen V. Fox, Ph.D.

Associate Professor of Psychology

Department of Psychology
Salisbury State College
Salisbury, Maryland 21801

Paper presented at the

Annual Meeting of the Eastern Educational Research Association,

Miami Beach, Florida, February 1988

TM012947

ABSTRACT

A comparison was made between scores and grades of students taking a Development and Learning course using either a Modified Mastery System or a Modified Norm-referenced System of Grading. Under the Mastery System students could take each unit exam up to three times to meet minimum or higher criteria levels. Under the M. Norm-reference System students took only one test per unit and were assigned grades by comparing their score on the unit (or course average) to the MEAN of the UPPER 10% of the distribution of scores (or averages). Although there was improvement in the M. Mastery group scores and a greater distribution of grades under the M. Norm-referenced System, repeated testing under the M. Mastery System did not yield significantly higher averages on unit tests or final course averages than found in the M. Norm-referenced group.

Although there may be teaching situations in which this is not the case, it has always been my objective to have as many students as possible, master as much information as possible, and evaluate that achievement validly. After teaching for approximately fifteen years I still find myself searching for an "ideal" grading system. Traditional criterion grading systems in which pre-determined criteria levels (i.e. 90% correct = A, 80% correct =B, etc.) are held to regardless of characteristics of a specific test, test items or student groups (difficulty of the test, wording of items, interests of students and resulting depth of coverage of material) are too rigid. "Curving," adding points to students' scores to compensate for those same characteristics is too arbitrary. Traditional norm-referenced grading systems, which assign grades based on a students' position within a distribution of scores are too competitive.

Mastery testing or PSI systems of instruction seemed to be the answer. It seems logical that if one wants to teach, one must accommodate to the different rates and styles of learning. Because this issue of testing and grading is most relevant to future teachers I used the Mastery system in my Development and Learning course, the first required course in our professional education program.

In the Mastery course students were given the following guidelines for testing and evaluation. A student's grade in the course was based on the average of three unit exams and one paper. Test scores and the final course average were converted to grades using the following criteria: 90% or higher correct (or final average) = A; 80-89% = B; 70-79% = C; 60-69% = D; below 60% correct (or final average) = F. Each exam included 40-60 multiple choice items that asked students to analyze, snythesize and apply concepts as well as remember definitions. Each student was given a maximum of three opportunities using parallel test forms to be evaluated on each unit. Only the highest grade on each unit was used in the calculation of the final course average. Students were encouraged to complete remedial activities such as writing outlines, attending tutorial sessions with the instructor, or seeking help at the reading/study skills centers available at the College, before taking a retest for a particular unit. Students returned all exams and answer sheets to the instructor following the exam itself and the review of the graded exam.

For many years I heard students using this Mastery system comment on how much they appreciated the opportunity to repeat exams in an effort to improve their scores and grades. I also

heard many comments excusing what I would consider less than the students best effforts on the exams, comments such as "I knew I could have done better but . . . I had a paper, test, assignment due in another class . . . and/or . . . I knew I could take this test over . . . ." Was I really helping a greater number of students master more information?

In January of 1987 I attended the 9th Annual National Institute on the Teaching of Psychology presented by the University of Illinois at Urbana-Champain* and was introduced to another grading system that seemed rational and humane. The same Development and Learning course was offered in the Spring and Fall of 1987 using the University of Illinois "Modified Norm-referenced" grading system. The course content and schedule were the same as in the Mastery course in the Spring and Fall of 1986. However, in 1987 students were given only one exam on each unit. Their course grade was based on the average of the three unit exams and one paper. The exams used in 1987 were the same exams used for the first attempt on each unit in 1986. Each exam score and the course average for each student was assigned a grade based on the following procedure:

*Presented by Dr. Frank Costin, Professor Emeritus of Psychology. University of Illinois at Urbana-Champain.

the upper 10% of the distribution of exam scores (or final averages) and the Mean of the upper 10% of scores or final averages) were determined. Grades were then assigned using the following scale: 95% of the mean (of upper 10%) = lower limit of A; 85% of the mean (of upper 10%) = lower limit of B; 75% of the mean (of upper 10%) = lower limit of C; 65% of the mean (of upper 10%) = lower limit of D. Using this method the percentage of A's, B's, C's, etc., is NOT determined in advance. A student's grade depends on what he/she achieves and not on a particular rank order. By using the mean of the upper 10% of scores as the comparison measure there is also consideration given to the specific characteristics of the test, items and/or group of students.

A comparison was made between students' test scores in the Modified Mastery course offered in the Spring and Fall of 1986 and students' test scores in this course as it was offered in the Spring and Fall of 1987 using the University of Illinois Modified Norm-referenced system.

The Modified Mastery and the Modified Norm-referenced groups were found to be equivalent in terms of group size (M. Mastery = 48, M. Norm-referenced = 52); class status

(approximately 34% sophomores and 49% juniors in each group); Majors (approximately 50% elementary education and 25% secondary or physical education majors in each group); Cumulative GPA's (approximately 2.8 in each group); and a CLOZE pre-reading score (approximately 60% correct average in each group) which is an informal assessment of how well a student's reading level matches that of the required text.

In the following data analysis only test scores were used. An analysis of the data by t-tests indicated that although there was a significant improvement in scores of M. Mastery students (Table I) on repeated exams there was no significant difference between the average scores achieved by the M. Mastery group after the first, second or third attempts on Units I, II or III, the highest score on each unit or the final course average and the M. Norm-referenced group (Table II). Although there was improvement in M. Mastery group scores, repeated testing did not yield higher average on unit tests or final course averages than found in the M. Norm-referenced group.

INSERT TABLES I AND II HERE

There were significantly more C's and B's given in the M. Mastery course. More A's and D's were given using the M. Norm-referenced system (Table III). For analysis only, the M. Norm-referenced scores were converted to grades using the criteria of 90% or higher = A; 80-89.9% = B; 70-79.9% = C; 60-69.9% = D; below 60% = F. When these assigned grades were compared to grades of students in the M. Mastery course no significant difference in frequency of grades was found (Table IV).

INSERT TABLES III AND IV HERE

After being given three opportunities to demonstrate competency in a content area we would expect at least the highest score achieved on each unit and final course average of the M. Mastery group to be significantly higher than the scores achieved by the M. Norm-referenced group. The fact that we found no significant differences between these groups leaves us with little justification for repeated testing and tutorial supports. Based on student comments I suspect that the advantage and disadvantage of the M. Mastery testing situation is that it lowered student anxiety (motivation?). While this lower anxiety

level may improve the performance of some students, many students seemed to use it as an opportunity to lower this course on their academic priority list. It is my impression that some students in the M. Mastery group did not make their best efforts in studying for the first form of each unit exam nor did they take full advantage of the repeated testing situation. Ultimately students' achievement under the two grading systems was not statistically different. While Mastery testing offers us a theoretically appealing system, the University of Illinois Modified Norm-referenced system offers us an acceptable alternative in Non-mastery learning situations.

TABLE I

COMPARISON OF UNIT AVERAGES WITHIN M. MASTERY GROUP

Number of Cases	Mean	Stand. Deviat.	Stand. Error	(Diff.) Mean	Stand. Devia.	Stand. Error	Corr.	2-Tail Prob.	T Val.	Degree of Free.	2-T Pr.
Unit I-Form A	72.9583	12.730	1.837								
48				-0.3333	7.717	1.114	0.836	0.000	-0.30	47	0.765
Unit I-Form B	73.2917	13.912	2.008								
Unit I-Form A	72.9583	12.730	1.837								
48				-4.3333	7.501	1.083	0.809	0.000	-4.00	47	0.000
Unit I-Form C	77.2917	9.876	1.425								
Unit I-Form A	72.9583	12.730	1.837								
48				-4.5208	7.400	1.068	0.817	0.000	-4.23	47	0.000
Unit I-Highest Grade	77.4792	9.442	1.363								
Unit II-Form A	74.5208	10.081	1.455								
48				-0.9167	3.836	0.554	0.928	0.000	-1.66	47	0.104
Unit II-Form B	75.4375	10.116	1.460								
Unit II-Form A	74.5208	10.081	1.455								
48				-1.1250	3.734	0.539	0.929	0.000	-2.09	47	0.042
Unit II-Form C	75.6458	9.692	1.399								
Unit II-Form A	74.5208	10.081	1.455								
48				-1.1250	3.734	0.539	0.929	0.000	-2.09	47	0.042
Unit II-Highest Grade	75.6458	9.692	1.399								

TABLE I - continued
COMPARISON OF UNIT AVERAGES WITHIN M. MASTERY GROUP

Number of Cases	Mean	Stand. Deviat.	Stand Error	(Diff.) Mean	Stand Devia.	Stand. Error	Corr.	2-Tail Prob.	T Val.	Degree of Free.	2-T Pr.
Unit III-Form A	72.3750	12.821	1.850								
48				-2.0000	6.760	0.976	0.854	0.000	-2.05	47	0.045
Unit III-Form B	74.3750	12.035	1.737								
Unit III-Form A	72.3750	12.821	1.850								
48				-3.5000	6.291	0.908	0.871	0.000	-3.85	47	0.000
Unit III-Form C	75.8750	10.965	1.583								
Unit III-Form A	72.3750	12.821	1.850								
48				-3.5000	6.291	0.908	0.871	0.000	-3.85	47	0.000
Unit III-Highest Grade	75.8750	10.965	1.583								

TABLE II

COMPARISON OF AVERAGE GRADES BETWEEN GROUPS
(Group I = M. Mastery; Group II = M. Norm-Referenced)

	No. Cases	No. Repeating Exam.	Mean.	S.D.	S.E.	F Value	2-Tail Prob.	Pooled Variance Estimate or Separate Variance Estimate T Value	D.F.	2-Tail Prob.
Unit I										
Form A										
GR I	48		72.96	12.73	1.84					
GR II	52		72.88	13.64	1.89	1.15	0.634	0.03	98	0.978
Form B										
GR I	48	27	73.29	13.91	2.00					
GR II	52		72.88	13.64	1.89	1.04	0.887	0.15	98	0.883
Form C										
GR I	48	10	77.29	9.88	1.43					
GR II	52		72.88	13.64	1.89	1.91	0.027	1.86	92.87	0.066
Highest Grade										
GR I	48		77.48	9.44	1.36					
GR II	52		72.88	13.64	1.89	2.09	0.012	1.97	91.07	0.052
Unit II										
Form A										
GR I	48		74.52	10.08	1.46					
GR II	52		73.81	10.81	1.50	1.15	0.630	0.34	98	0.734
Form B										
GR I	48	7	75.44	10.12	1.46					
GR II	52		73.81	10.81	1.50	1.14	0.65	0.78	98	0.439
Form C										
GR I	48	1	75.65	9.69	1.40					
GR II	52		73.81	10.81	1.50	1.24	0.45	0.89	98	0.374
Highest Grade										
GR I	48		75.65	9.69	1.40					
GR II	52		73.81	10.81	1.50	1.24	0.45	0.89	98	0.374

TABLE II - continued
COMPARISON OF AVERAGE GRADES BETWEEN GROUPS
(Group I = M. Mastery; Group II = M. Norm-Referenced)

	No. Cases	No. Repeating Exam.	Mean	S.D.	S.E.	F Value	2-Tail Prob.	Pooled Variance Estimate or Separate Variance Estimate T Value	D.F.	2-Tail Prob.
Unit III										
Form A										
GR I	48		72.38	12.82	1.85					
GR II	51		75.37	12.61	1.77	1.03	0.91	-1.17	97	0.244
Form B										
GR I	48	21	74.38	12.04	1.74					
GR II	51		75.37	12.61	1.77	1.10	0.75	-0.40	97	0.688
Form C										
GR I	48	2	75.88	10.97	1.58					
GR II	51		75.37	12.61	1.77	1.32	0.34	0.21	97	0.833
Highest Grade										
GR I	48		75.88	10.97	1.58					
GR II	51		75.37	12.61	1.77	1.32	0.34	0.21	97	0.833
Course Average										
GR I	48		76.34	8.93	1.29					
GR II	52		73.54	11.99	1.66	1.81	0.04	1.33	93.87	0.187

TABLE III

COMPARISON OF COURSE GRADES

As Assigned by Different Methods

Grade	Count	M. Mastery	M. Norm Ref.	Row Total
F	0	2	3	5
D	1	3	13	16
C	2	23	14	37
B	3	18	13	31
A	4	2	9	11
	Column Total	48	52	

Chi-Square	D.F.	Significance	Min E.F	Cells with E.F. < 5
13.76221	4	0.0081	2.400	2 of 10 (20.0%)

Number of Missing Observations = 0

TABLE IV

COURSE AVERAGE MATHEMATICALLY ASSIGNED GRADE

Grade	Count	M. Mastery	M. Norm Ref	Row Total
F	0	2	5	7
D	1	3	12	15
C	2	26	17	43
B	3	16	15	31
A	4	1	3	4
	Column Total	48	52	

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F. < 5
9.45682	4	0.0506	1.920	4 of 10 (40.0%)

Number of Missing Observations = 0